ABSTRACT OF THE DISCLOSURE

A computer implemented method and program for analyzing statistical an original data set having a first size, dimension and distribution. Multiple random data sets are generated, each having a second size, dimension and distribution related to the first size, dimension and distribution of the original data set. Numerical values of test statistics corresponding to the random data sets are calculated in accordance with a predetermined test statistic formula. A relationship between the numerical values corresponding to the random data sets and the numerical value of the test statistic corresponding to the random data set, calculated in accordance with the test statistic formula, is determined. It is determined that the original data set includes at least one factor not based on chance when the relationship indicates that the numerical value of the original test statistic is not within a range of the numerical values corresponding to the random data sets.